Reviewer's report

**Title:** Gene expression profiles of breast biopsies from healthy women identify a group with claudin-low features

**Version:** 1  **Date:** 14 September 2011

**Reviewer:** Robert B Clarke

**Reviewer's report:**

The paper has identified a new feature of some breast biopsies from ostensibly normal breast present in both healthy women and in benign breast tissue from women with breast cancer. This sub-group of samples termed ‘cluster 1’ have a gene expression profile that has claudin-low, mesenchymal (EMT-like) and possibly stem cell-like features. This is extremely interesting but largely intriguing since in the discovery and validation sample cohorts used there is no adjacent tissue available for histological and pathological examination. There is some correlation to nulliparity, family history but in this small cohort there are 4 breast cancers associated with the ‘cluster 2’ type gene expression and none in ‘cluster 1’. However, the numbers and follow-up are not sufficient for any conclusions to be drawn in this regard.

Overall, the analysis is very well performed and the gene expression differences between the clusters are robust in the 2 cohorts of women examined. Ideally, a third cohort should be added to investigate the underlying histology/pathology.

However, I believe the finding and its implications should be published in order to open up this novel line of enquiry as to the origins and risks associated with this claudin-low gene expression cluster. At present, the current report is a preliminary but important observation.

Thus, discretionary revision that would strengthen the finding and patently increase the impact would be the addition of the third cohort that describes the histological features underlying ‘cluster 1’.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.